# Capítulo 2

# ESTÁNDARES Y EVALUACIÓN

*Gilbert A. Valverde*

## Introducción

Los estándares y la evaluación son dos instrumentos importantes del currículo prescrito de un sistema educativo. El currículo prescrito[1] se ha vuelto un tema prominente en la política educativa a nivel global desde la segunda mitad del siglo XX. Al igual que en América Latina, muchos de los sistemas educativos del mundo experimentaron un importante giro en sus prioridades en ese período. Mientras que en tiempos anteriores la política educativa había enfocado principalmente la inversión en escuelas y otros recursos materiales similares, como parte de una agenda destinada a incrementar o asegurar el acceso universal a la educación, las décadas de 1980 y 1990 trajeron un énfasis mayor sobre el conocimiento conceptual, procedimental y otros objetivos académicos que debían ser alcanzados por los estudiantes en la educación primaria y secundaria. Como consecuencia de esto, surge un renovado interés por el currículo prescrito como una herramienta importante para la política educativa. El movimiento hacia la creación, en algunos casos, o la redefinición en otros, de sistemas nacionales de pruebas, estándares e instrumentos similares refleja este énfasis en la calidad del contenido de la escolarización. Estas ideas surgen a medida que los diseñadores de políticas y administradores de los sistemas educativos favorecen la formulación de currículos nacionales

---

1 Aquí se empleará el término «currículo prescrito» o «currículo intencional» siguiendo el modelo común en los estudios como el Tercer Estudio Mundial de Matemáticas y Ciencias (TIMSS) –referido a las prescripciones curriculares a nivel macro del sistema educativo. El modelo contempla además dos niveles: el currículo ejecutado o implementado (las experiencias educativas en la escuela y el aula), y el currículo logrado (los conocimientos, habilidades y actitudes que adquieren los y las estudiantes como consecuencia de su escolarización (Valverde 1997).

y una variedad de instrumentos de implementación con el fin de asegurar la ejecución y el logro de contenidos y habilidades disciplinarios considerados socialmente valiosos.

La mayoría de las nuevas políticas curriculares estipulan la adquisición de habilidades cognoscitivas de orden superior por parte de todos los estudiantes, y estas prescripciones suelen inspirarse en concepciones acerca del tipo y la cantidad de conocimientos y habilidades que se perciben que son necesarios para ejercer plenamente los derechos humanos, ciudadanos y económicos en sus respectivas naciones.

En América Latina, igualmente existe discusión acerca de la importancia de la calidad del contenido de la escolarización. Estas discusiones han promovido la evaluación y, más recientemente, los estándares, como instrumentos importantes a considerar. No obstante, hay indicios de que persisten importantes problemas de calidad. Pruebas nacionales, regionales e internacionales han documentado la pobreza de los resultados de la escolarización en América Latina, y sugieren que el repertorio actual de políticas orientadas a mejorar la calidad es insuficiente.

La discusión misma en torno a la calidad educativa en Latinoamérica suele ser muy vaga y limitada, y temas urgentes se encuentran a menudo fuera del temario de la discusión. Por ejemplo: el tema del rigor académico, que se expresa en el nivel del desafío cognoscitivo planteado en las metas curriculares en América Latina, no es un tópico comúnmente abordado en las políticas curriculares. Rara vez este tema recibe mención y, sin embargo, ciertamente el perfil de las expectativas curriculares en áreas como la matemática, por ejemplo, difícilmente puede caracterizarse como de alto rigor en la región.

Tampoco se da mucha discusión (más allá de elementos imprecisos o superficiales) acerca del tema de las habilidades para la vida ciudadana o económica –una perspectiva diferente a lo que algunos podrían objetar como el sesgo «academicista», que supone una perspectiva centrada en nociones de rigor disciplinario.

Las discusiones de política en la región suelen omitir también la discusión de la relación entre el currículo prescrito o intencional y la evaluación nacional. La noción de que la evaluación nacional operacionaliza[2] el currículo prescrito para propósitos de medición y monitoreo o propósitos de responzabilización[3]

---

[2] A lo largo de este capítulo se utiliza la palabra operacionalización, tomando su significado derivado de las ciencias sociales empíricas: se trata de definiciones que especifican las operaciones que se deben realizar para confirmar empíricamente (mediante la observación o medición) la presencia de un constructo, una meta, un estándar, un objetivo, etc. en el mundo social.

[3] Se prefiere usar el término «responsabilización» como alternativa a la frase «rendición de cuentas» con la que se traduce comúnmente la palabra *accountability*. «Rendición de cuentas»

(*accountability*) existe en algunos casos, pero no se ha derivado de esta idea una consideración específica de las articulaciones concretas entre la política con respecto al currículo prescrito, y de su instrumentalización técnica en las evaluaciones nacionales. Se nota en la región una especial inatención a los aspectos de construcción de consensos que supone la introducción de nuevos instrumentos políticos, como los estándares o las evaluaciones con propósito de rendición de cuentas, necesarios para que estos se tornen en efectivos instrumentos para lograr calidad.

Es en este contexto que el objetivo de este capítulo intenta abordar el tema de los estándares y de la evaluación.[4] Estos tópicos se tratan en tanto herramientas de política curricular que pueden estructurar y adelantar conversaciones públicas acerca de los objetivos de la escolarización en América Latina y que, desafortunadamente, han sido postergados en muchas de nuestras naciones. En este capítulo se revisan elementos de la experiencia a nivel mundial en relación con estos temas, usando ejemplos e ilustraciones tomados también de la región latinoamericana. Se parte de una discusión de la política curricular examinando su estado actual en América Latina y explorando tendencias a nivel internacional. Asimismo, se revisan los tipos de currículos prescritos existentes en Latinoamérica y se discute cómo estos se comparan con las políticas educativas de países que han demostrado éxito en lograr niveles competitivos de desempeño promedio de sus estudiantes en pruebas internacionales –lo que podría llamarse «estándares de nivel mundial».

Siguiendo esta discusión, se aborda una consideración respecto de la evaluación en América Latina, teniendo en cuenta la cuantiosa experiencia regional que hay en esta área. Por otra parte, se hará una revisión del balance de esa experiencia, procurando identificar el estado de avance actual y los desafíos que quedan por resolver. Finalmente, se examina la articulación de los estándares y la evaluación dentro de una estructura de políticas curriculares. Respecto a la discusión de esta estructura de políticas curriculares, se pone el énfasis en cómo

se considera que indica más bien fiscalización, cuando en realidad el término supone más certeramente un sistema de responsabilidades recíprocas.

[4] El autor de este capítulo es miembro del Grupo de Trabajo sobre Evaluación y Estándares (GTEE) del PREAL. Gran parte del material descriptivo de este capítulo deriva de distintos trabajos llevados a cabo y auspiciados por el GTEE (Ferrer 2006; Ravela et al. 2001). En cuanto a investigaciones internacionales comparativas, en su mayoría pertenecen al autor o al equipo de investigadores del centro de Estados Unidos para el Tercer Estudio Mundial sobre Matemática y Ciencias, del cual formó parte. Estos se complementan con entrevistas realizadas por el propio el autor en diversas visitas a las unidades de evaluación, currículo y equipos encargados de estándares, además de conversaciones sostenidas con distintos actores de decenas de sistemas educativos de Europa, Asia, y América Latina a lo largo de los últimos 15 años.

se podrían organizar conversaciones más fértiles acerca de metas y prioridades relacionadas con la calidad educativa. Se procura al mismo tiempo aclarar cómo una estructura de políticas curriculares, en la que los estándares y la evaluación se articulan en una arquitectura completa de instrumentos políticos convergentes, puede ayudar a favorecer la creación de mejores oportunidades para aprender en los sistemas educativos de la región.

## I. Claridad de expectativas

La investigación educativa y la experiencia política indican que lograr alta calidad en los resultados de la escolarización está vinculado con la articulación de una visión clara de los objetivos pedagógicos que especifiquen en términos operacionales una visión de la calidad que sea rigurosa, desafiante y centrada en las disciplinas académicas (Benavot 1992; Benavot et al. 1991; McKnight y Valverde 1999; Schmidt et al. 2001; Valverde y Schmidt 2000; Westbury y Hsu 1996). Un principio consecuente de equidad indica que todos los estudiantes de un país deberían tener las mismas oportunidades para lograr expectativas de este tipo, para lo cual es importante especificarlas y comunicarlas a los actores principales del sistema en forma clara, operacional y ejemplificada.

Existen distintos documentos en diversos países de la región que prescriben lo que deben aprender los estudiantes en sus escuelas y colegios. Como resultado de procesos de reforma inspirados en concepciones de calidad, los documentos que comunican estas metas proponen una nueva dirección para la educación y en ellas se intenta plasmar una visión renovadora y moderna acerca de las áreas curriculares. Sin embargo, una mirada crítica sugiere que estas formulaciones deben ser refinadas para corregir problemas de ambigüedad, de dispersión y contradicción, y de rigor académico.

La ambigüedad se observa en las formulaciones poco específicas y vagas que se encuentran a menudo presentes en los documentos curriculares de base en nuestros países. Al leerlos, el lenguaje y la composición no permiten, en muchos casos, discernir cómo se puede reconocer en el salón de clase el objetivo referido. Este problema conduce a una serie de reformulaciones del mismo mensaje acerca de los objetivos (en planes y programas de estudio, por ejemplo) que, en su intento por resolver la ambigüedad, más bien la agravan, ya que proponen nuevas definiciones, interpretan de manera inconsistente los objetivos originales y, a veces, descartan algunos de sus elementos centrales.

Ciertamente, muchas de las políticas curriculares de la región presentan, en un mismo currículo, objetivos con distintos niveles de imprecisión o especificidad. En el caso de la República Dominicana, el currículo hoy en día vigente

fue publicado en 1995. Dos ejemplos de los extremos de variedad en la formulación de objetivos se pueden ver para el caso de Ciencias Naturales y Lenguaje y Comunicación en el siguiente recuadro.

RECUADRO Nº1
EJEMPLOS DE PROPÓSITOS CURRICULARES EN EL CURRÍCULO DE LA REPÚBLICA DOMINICANA
(Secretaría de Estado de Educación, 1995)

**En Ciencias de la Naturaleza (4º Grado):**
«Valorar la importancia de los recursos naturales y protegerlos.» (pág. 237)

**En Lengua Española (4º Grado):**
«Dialogar con elementos de planificación sencilla: intercambios, reuniones de trabajos en grupos, con amigos o compañeros del entorno social cercano de la escuela a partir de un tema, un propósito, un interés previamente definidos.» (pág. 120)

En este ejemplo, las metas son muy distintas. No solamente porque a propósito el currículo especifica distintos tipos de contenido, ya sea conceptuales, procedimentales o actitudinales. Si bien en el currículo los tres tipos de contenido comparten una característica: la de proponer una visión motivadora, una identificación de aspectos generales del sentido de la educación en ciencias y lenguaje para la educación dominicana; sin embargo, los objetivos en ciencias contienen vocablos que poseen tantos potenciales referentes específicos y connotaciones («valorar», «importancia»), que resulta imposible que las metas sean usadas como referentes concretos para determinar si un niño o una niña ha alcanzado el objetivo. Inclusive la parte más crítica de la meta, la idea de «protección» de los recursos naturales, alude a un contexto externo al aula, en el que los actores del sistema educativo no tendrán la oportunidad de hacer observaciones acerca de los logros individuales de los y las estudiantes. En el caso del objetivo de Lengua Española, a primera vista parece presentar la meta en una forma más específica. Sin embargo, también se caracteriza por su vaguedad: ¿cuáles serían los elementos de planificación sencilla? ¿Cuáles serían las evidencias de que se ha llevado a cabo un diálogo a partir de un propósito o interés previamente definido? Por otro lado, la meta sugiere una pluralidad de elementos de planificación y enumera tres tipos de diálogo con distintos niveles de especificidad (por ejemplo «intercambios» versus «trabajos en grupo»), y también múltiples tipos de interlocutores. Siendo así, ¿cómo se podría

confirmar si el objetivo ha sido logrado por parte de un alumno o una alumna en particular? El tipo de problema de vaguedad y ambigüedad que ilustra este pequeño ejemplo es común en las políticas curriculares de toda la región. Como se aprecia en la ilustración, los tres tipos de meta comparten otra característica: la vaguedad y la falta de definiciones operacionales.

La dispersión se observa precisamente en la multiplicidad de formulaciones curriculares que existen en un mismo país o sistema educativo, cada una intentando desglosar objetivos sistémicos en las distintas áreas del currículo de manera diversa, conduciendo a un gran número de mensajes curriculares discordantes acerca de los objetivos que los docentes deben de perseguir en el salón de clases.

La contradicción se observa en el hecho de que las diversas formulaciones y reformulaciones de los objetivos curriculares frecuentemente no se encuentran alineadas,[5] magnificándose así los efectos de la ambigüedad y la dispersión.

Un buen ejemplo de esto se presentaba en Perú en el año 2002. La relación entre la Estructura Curricular Básica (ECB) del primer ciclo de educación primaria en ese país (Dirección Nacional de Educación Inicial y Primaria 2000) y el documento ministerial que proponía logros básicos e indicadores (Ministerio de Educación 1997) para complementar la ECB era contradictoria. En el primero se identificaban tres competencias para el área de desarrollo[6] «ciencia y ambiente»,[7] mientras que en el segundo estas tres metas fueron reformuladas en un lenguaje nuevo –pero igualmente genérico– bajo el nombre de «logros básicos».[8] Los «indicadores de logro por grado» reformulaban de forma similar, en lenguaje poco operacional y sin ejemplos, las «capacidades y actitudes» de la ECB. Este patrón de desarrollo curricular, que consiste en incluir distintas «capas» de mensajes reformulados, utilizando vocabulario distinto y sin lenguaje operacional, se evidencia a lo largo de los distintos documentos curriculares y constituye una característica que comparten los instrumentos de política curricular de muchos países de América Latina. El resultado es un conjunto de

[5] Es decir, sus especificaciones de expectativas no son congruentes al no compartir definiciones comunes.

[6] La diversidad de términos usados en América Latina en los distintos instrumentos de política curricular también dificulta la discusión. Un «área de desarrollo» en el currículo peruano se refiere a una asignatura o área temática del currículo –lo que en otros países se podría llamar «ramo», «eje», «curso», etc.

[7] i) Conocimiento de su cuerpo y conservación de la salud en armonía con su ambiente, ii) conocimiento y conservación del ambiente y iii) intervención humana en el ambiente.

[8] i) Maneja información y procesos para el conocimiento de su cuerpo y su ambiente, ii) reflexiona e investiga problemas físicos y sus interrelaciones con el ambiente, y iii) aplica conocimientos científicos y tecnológicos para la resolución de problemas prácticos y valora la intervención del hombre en el ambiente.

mensajes ambiguos y a menudo contradictorios.[9] A veces inclusive la multiplicidad de mensajes no es consecuencia accidental de la acumulación de distintos documentos que pretenden desglosar o explicar un mismo currículo, como en este caso. De hecho, hay países en la región en los que subsisten al mismo tiempo, para una misma población estudiantil, distintos currículos prescritos. En la educación pública de Uruguay coexisten, en el momento que escribimos este capítulo, no menos de tres planes de estudio en Matemática en secundaria, cada uno con diferente difusión en cuanto al número y la ubicación geográfica de los centros educativos en donde se aplica. Inclusive, las instituciones de educación secundaria privada tienen derecho a optar por algunos programas de estudios previos a los oficiales vigentes.

Una revisión de docenas de currículos y estándares[10] de la región arrojan un hallazgo importante: es notable la ausencia, entre los instrumentos curriculares de América Latina, de un abordaje claro y explícito de la importancia del rigor en el currículo. Es poco frecuente, incluso, que los objetivos curriculares mismos aborden tópicos o temas académicamente rigurosos y que requieran de funciones cognitivas complejas.[11] Esto distingue a la actual política curricular

---

[9] Como ejemplo, la competencia «Conocimiento de su cuerpo y conservación de su salud en armonía con su ambiente» (ECB-1° Ciclo), es desglosada en una serie de capacidades y actitudes relacionadas con salud, higiene y nutrición. Sin embargo, esta misma competencia, ahora identificada con el logro básico «maneja información y procesos para el conocimiento de su cuerpo y su ambiente», se desglosa en indicadores para primer y segundo grados que incluyen un número importante de estos relacionados con la anatomía y morfología animal –aspectos no incluidos en esta competencia en la ECB.

[10] Existen tanto estándares como otros instrumentos de política curricular en la región. No son idénticos, pero sí conforman ambos instrumentos del currículo prescrito. En la penúltima sección de este capitulo, se definen los estándares y los otros instrumentos de política curricular para señalar sus diferencias. En esta parte del capítulo, se tratan todos de forma semejante, en cuanto instrumentos del currículo prescrito que existen para formular objetivos para la escolarización.

[11] Evidentemente, la noción de *rigor* depende de la disciplina académica y del grado o nivel educativo al que corresponde. Metas rigurosas en Matemática, por ejemplo, incluyen aquellas que en la educación secundaria abordan temas centrales de geometría como posición, visualización, forma, simetría, congruencia y semejanza. Temas de proporcionalidad y también funciones, relaciones y ecuaciones, culminando por lo general con temas introductorios de análisis elemental como procesos infinitos (sucesiones aritméticas y geométricas, series aritméticas y geométricas, teorema del binomio, otras sucesiones y series, límites y convergencia de series, límites y convergencia de funciones, continuidad) y cambio (crecimiento y decrecimiento, diferenciación, integración, ecuaciones diferenciales, diferenciación parcial). Suele ocurrir que estas metas buscan desarrollar en los estudiantes habilidades para representar, reconocer equivalencias y realizar procedimientos rutinarios –*buscando al mismo tiempo* desarrollar habilidades más complejas, como por ejemplo investigar y resolver problemas o razonar matemáticamente (mediante la generalización, establecimiento de conjeturas, justificación y demostración, y axiomatización).

de América Latina de gran parte de los países del mundo, cuya preocupación apunta a reformar la educación según nuevos criterios de calidad.

Se podría argumentar, por lo contrario, que la discusión del rigor es menos importante que el rigor mismo. Es decir, importa más que las políticas curriculares propongan metas desafiantes para alcanzar niveles de desempeño que evidencian habilidades cognoscitivas complejas en las distintas disciplinas del currículo, que contengan una simple «declaración de principios» que discuta la importancia del rigor. Sin embargo, al estudiar los currículos de América Latina, la ausencia de una discusión acerca del rigor disciplinario parece ser un indicador de que este no ha constituido un criterio de importancia en la elaboración de tales políticas. La evidencia al respecto varía. Primero, casi todos los documentos curriculares contienen presentaciones acerca de la filosofía pedagógica-didáctica que fundamenta la propuesta curricular, y en estas discusiones tiene un lugar prominente la discusión de teorías de construcción psicosocial del conocimiento, perspectivas pedagógicas centradas en el alumno y otros similares. Las discusiones se presentan a menudo como declaraciones de principio sin ser acompañadas por detalles específicos de cómo estos se aplican a los propósitos concretos de las distintas áreas curriculares del currículo. Tampoco es frecuente que se preocupen por una fundamentación empírica o sustentada en evidencias para la elaboración de las propuestas curriculares, reflejando quizá la inexistencia de información apropiada derivada de evaluaciones o investigaciones empíricas nacionales sobre el aprendizaje de niñas y niños en el sistema escolar, o porque las tradiciones intelectuales con respecto al currículo en la región han concedido históricamente poca importancia a la investigación empírica. Por tanto parece prevalecer la opinión de que los currículos deben afianzarse en concepciones filosóficas globales y no en consideraciones de rigor a partir de las disciplinas.

Una explicación de lo anterior, probablemente aplicable a los currículos más recientes, es que se evitan concepciones de rigor académico en forma deliberada. La razón podría radicar en que tales concepciones se asocian a un enfoque tradicional academicista que es percibido como teórico y estéril con respecto a las exigencias del mundo contemporáneo. Otra posibilidad es que muchos de los proponentes de las nuevas perspectivas curriculares consideran que estas perspectivas son incompatibles con prescripciones curriculares de todo tipo y por ende, cuando se ven forzados a escribirlas, prefieren preparar documentos que privilegian la discusión teórica-filosófica de las teorías pedagógicas que inspiran la nueva perspectiva curricular, pero que son intencionadamente imprecisos y generales al discutir el currículo de las distintas asignaturas escolares. Inclusive, mucha de la tradición intelectual de la región podría militar en contra de la formulación precisa de objetivos.

La teoría curricular tradicional en la región, inspirada en la «nueva sociología» del currículo de la década de 1970 (Bernstein 1973) o en la teoría de la dependencia (Cardoso y Faletto 1980; Devon 1983; Sica y Prechel 1981) surgida en gran medida de algunos economistas políticos e historiadores de economía de América Latina y África, principalmente, centra gran parte de su atención en la estructura de relaciones económicas y políticas –entre países, distintos grupos dentro de un mismo país e individuos– que se promueven o reproducen en los currículos. Como gran parte de la discusión acerca de la calidad se ha tornado «global», con interés creciente en la comparación internacional de objetivos y la participación en enormes pruebas internacionales de logros con sus tablas de posiciones comparativas de los logros de estudiantes de distintos países, posiblemente se vuelve difícil promover la discusión bajo esta concepción en muchos círculos intelectuales, cuyo frecuente escepticismo o inclusive hostilidad a la globalización es una manifestación actualizada de esta importante tradición intelectual.

Pero Latinoamérica, al sacrificar el rigor como uno de los criterios a tomar en cuenta, junto con perspectivas filosóficas orientadoras, hace caso omiso a la evidencia que se ha venido acumulando con respecto al hecho de que el mundo del trabajo y el mundo de la participación ciudadana en las democracias han experimentado importantes incrementos en el nivel de su demanda cognoscitiva. El resultado son políticas curriculares en las cuales el currículo está débilmente articulado con estas demandas cognoscitivas del trabajo y la vida ciudadana.[12] Prevalecen prescripciones curriculares caracterizadas por recomendar temarios superficiales, con la ausencia de muchos de los grandes temas o teorías que les dan sentido a las disciplinas.[13] Asimismo, se descarta la considerable evidencia empírica de que una de las maneras más eficaces de promover oportunidades educativas que aumenten la calidad educativa es precisamente formular instrumentos que señalen de manera clara objetivos curriculares rigurosos, desafiantes y significativos.

---

[12] Gran cantidad de trabajo reciente ha documentado la creciente demanda por habilidades y conocimiento que exige la participación ciudadana eficaz y el mundo del trabajo –esto incluye el mundo del trabajo manual (Levy y Murnane 2005; Schneider 2005; Torraco 2002). Por otro lado, un estudio en 1999 de 40 documentos curriculares de 14 países latinoamericanos documentó exhaustivamente la prevalencia de tratamientos de temarios superficiales y poca articulación temática en ellos (Ferrer et al. 2006).

[13] Entre las grandes teorías ausentes en muchos currículos de educación secundaria en Ciencias Naturales, se encuentran la teoría de la evolución y la teoría de la relatividad. Llama la atención, por ejemplo, que en muchos países latinoamericanos, un estudiante puede graduarse de la educación secundaria sin haber estudiado con algún nivel de profundidad la física posterior a Newton. Esto equivale a ignorar casi 300 años de importantes adelantos.

Estas características de vaguedad, ambigüedad, dispersión, contradicción y falta de rigor representan un desafío importante. Señalan debilidades centrales en las políticas referentes a la calidad de los resultados de la escolarización. Muchos países del mundo han hecho esfuerzos por construir instrumentos políticos que pretenden evitar muchas de estas dificultades. Por lo tanto, existen distintas opciones para fijar expectativas claras y de alto rigor y exigencia académica respecto de lo que deben aprender los y las estudiantes que pueden servir de ejemplo para la región.

## 1. Expectativas establecidas a partir del currículo

Un procedimiento para fijar expectativas claras se genera a partir de un sistema de metas curriculares que actúe como instrumento fundamental para alinear recursos, procesos e instrumentos de política educativa. Este tipo de sistema ofrece la posibilidad de que en él se establezcan criterios con relación al *progreso* de los estudiantes hacia las expectativas establecidas. Esa noción de progreso está a menudo del todo ausente, o poco desarrollada, en los sistemas de expectativas establecidas a partir de pruebas. Este método también permite incluir dimensiones axiológicas que no serían tan fáciles de fijar a partir del establecimiento de expectativas exclusivamente centradas en un sistema de evaluación.

Un ejemplo de un sistema de expectativas a partir del currículo es el de Japón. Existe en ese país un sistema de expectativas que forman un conjunto de instrumentos escritos a nivel del Ministerio Nacional, de las instancias en las distintas prefecturas, escuelas, etc., en donde, por ejemplo, la «Guía para la escuela básica», escrita por el ministerio, fija metas generales llamadas «objetivos», a las cuales se asocian desgloses llamados «de contenidos», que a su vez proporcionan información específica acerca de procesos e informaciones que se pretende que los estudiantes manejen. Un ejemplo, también tomado del área de las ciencias, se encuentra a continuación. En el caso de Japón, el objetivo general constituye la meta genérica, y el progreso o avance hacia el logro del objetivo se explicita en objetivos específicos por grado.

Es importante subrayar que las expectativas en Japón forman parte de un sistema –existen distintos instrumentos, escritos para distintas audiencias y propósitos. Su diferencia con el conjunto de mensajes curriculares en muchos países de América Latina radica en que las definiciones fundamentales de las expectativas son comunes –garantizando la congruencia o alineamiento entre instrumentos. De este modo, los libros de texto, los programas de capacitación y formación, y las pruebas escritas de la evaluación sistémica apoyan, juntos, un conjunto de expectativas comunes.

RECUADRO Nº 2
GUÍA PARA LA ESCUELA BÁSICA, CIENCIAS, JAPÓN
(MONBUSHO-Ministry of Education, Science and Culture, 1995)

1. Objetivo

   Lograr familiaridad con la naturaleza y llevar a cabo observaciones, experimentos y otros estudios – y mediante estos cultivar la habilidad para resolver problemas y propiciar un genuino afecto por la naturaleza. Estimular la comprensión de fenómenos naturales y propiciar puntos de vista y el pensamiento científico.

2. Objetivos y contenidos por grado.

**Tercer Grado**

1. Objetivos

   Examinar y comparar plantas y animales familiares –y también el cuerpo humano, y estudiar con interés problemas que se descubren– cultivando así actitudes afectuosas hacia las criaturas vivientes y ampliar puntos de vista y pensamientos acerca de sus características físicas y los principios de su crecimiento.

   (Siguen otros objetivos por grado...)

2. Contenidos

   A. Criaturas vivientes y sus medios

      (1) Buscar, recolectar y cultivar plantas familiares, a fin de desarrollar la capacidad de examinar procesos de crecimiento y características físicas, tales como:

         a) las plantas crecen siguiendo ciertas secuencias, y algunas crecen a partir de raíces y/o tallos, otras de semillas;
         b) el cuerpo de la planta se conforma de raíces, tallo, hojas, etc., y la composición física es distintiva para cada especie...

   (Siguen especificaciones de otros contenidos...)

Otra alternativa es el modelo que se sigue en algunos estados australianos o en el estado de Nueva York en Estados Unidos, donde la piedra angular del sistema de expectativas es un documento que establece estándares de aprendizaje y que incluye una especificación de distintas etapas o estadios de avance,

que marcan hitos importantes en el logro del objetivo.[14] Un inconveniente importante de este tipo de especificación de metas es que agrega un documento más a la cantidad de instrumentos que deben manejar los diversos actores del sistema. Sin un esfuerzo técnico riguroso, esto aumenta la probabilidad de que el «ambiente curricular» en el que se desenvuelven algunos actores principales del sistema (magisterio, estudiantes, autores y editores de libros de texto, etc.) sea aún más confuso. Otro factor que suele contribuir a la confusión es que a menudo estos objetivos se fijan ignorando el problema de la formación y capacitación de docentes, y cuando esto sucede, frecuentemente los documentos que resultan tienen un lenguaje formal, poco significativo para el magisterio, y no se encuentran acompañados por distintos mecanismos que apoyen los cambios en la pedagogía de acuerdo con los objetivos establecidos. A veces tampoco se vinculan las metas con el sistema de monitoreo y evaluación, y cuando esto pasa también el lenguaje es a menudo excesivamente formal y genérico –poco operacional. Tanto el magisterio como el sistema de monitoreo y evaluación de un país tienen experiencia relevante en la operacionalización del currículo que, al no aprovecharse en el momento de la formulación de metas, estas últimas resultan casi siempre en expresiones formales sin posibilidades de implementación.

En América Latina existe el interés, en algunos países, de buscar instrumentos de este tipo –como por ejemplo, los estándares. Sin embargo, la exigüidad de investigación empírica y discusión acerca del currículo prescrito «abierto» (en contraposición a lo que es llamado en la «nueva» sociología del currículo el «currículo oculto»), aunque importante en el estudio de la dolorosa historia de la repartición inequitativa de oportunidades educativas en la región, provee muy pocas luces con respecto a la pregunta acerca de cuáles serían las características de un currículo prescrito tendiente a mejorar los resultados de la escolarización.

A nivel mundial, un aspecto importante de la discusión acerca del currículo ha sido su carácter comparativo e internacional. Los expertos en currículo, las asociaciones profesionales y gremiales, y los diseñadores de políticas se preocupan por cómo los estándares definidos en sus países se comparan con los de otros países, especialmente aquellos países que consideran importantes referentes culturales, de calidad, o socios económicos. Estos países consideran de suma importancia, por ejemplo, determinar qué tipo de expectativas en política curricular caracterizan a los países que han probado lograr altos niveles de

---

14 En realidad, el modelo australiano influyó en el modelo de Nueva York. Actualmente, el Ministerio de Educación chileno está escribiendo objetivos curriculares orientados también por el modelo australiano. Presentamos aquí un ejemplo tomado de Australia en páginas posteriores.

rendimiento promedio de sus estudiantes en las pruebas internacionales comparativas a gran escala que se aplican regularmente entre los países desarrollados de Europa y Asia –y en las que comienza a aumentar el número de países participantes latinoamericanos.

América Latina, salvo notables e importantes excepciones, ha estado en gran medida ausente de esta discusión. Similar a la ausencia de una discusión del rigor, está ausente también una preocupación acerca de cómo se comparan las expectativas curriculares de sus países con las expectativas de los países que han logrado resultados sobresalientes en la escolarización de sus estudiantes a nivel mundial. Entre los países excepcionales en este sentido –que han incorporado la comparación de sus expectativas con expectativas de «nivel mundial» en el diseño de estándares o similares– se encuentran los recientes esfuerzos por desarrollar estándares en Chile, Guatemala y Honduras. En los tres países, la comparación internacional ha jugado un papel en las técnicas recientes que han utilizado para desarrollar estándares y son señal de un interés por aprender de otros países acerca de cómo promover oportunidades educativas con demostrable impacto sobre la calidad educativa, iniciativa que recién comienza en la región.

Desde mediados de la década de 1990 se ha venido explorando a nivel internacional la relación de las políticas curriculares con respecto a las oportunidades educativas que promueven o restringen el aprendizaje, y los logros académicos que resultan de estas oportunidades.[15] Enfocándose en caracterizaciones exhaustivas del contenido y expectativas de rendimiento en los estándares y libros de texto del mundo, algunos de estos estudios han arrojado informaciones con importantes implicancias para el desarrollo de la política curricular.

Estos hallazgos apuntan a una variedad de rasgos comunes a los países de alto rendimiento promedio que no comparten los países de rendimientos inferiores. Forman parte de lo que parece ser condiciones necesarias, pero no suficientes, para la realización de más altos niveles de logro escolar entre estudiantes.

Muchos países de bajo rendimiento tienen currículos que promueven la cobertura de largas listas de contenidos y habilidades en cada grado. En los países de alto rendimiento, sin embargo, es más común encontrar que estipulen la enseñanza y aprendizaje de un número más limitado de objetivos. Estos objetivos comúnmente debían explorarse a mayor profundidad que en el caso de los países de bajo rendimiento, y su finalidad es que los y las estudiantes deben dominarlos antes de pasar a nuevos objetivos en grados subsiguientes. Un currículo

---

[15] Esta sección resume hallazgos del Tercer Estudio Mundial de Matemáticas y Ciencias (TIMSS) que incluyó el primer estudio comparativo mundial exhaustivo del currículo prescrito (McKnight y Valverde 1999; Schmidt et al. 1999; Schmidt et al. 2001; Schmidt, McKnight, Raizen et al. 1997; Schmidt, McKnight, Valverde et al. 1997; Schmidt, Raizen, Britton et al. 1997; Stigler 1997; Valverde 2000; Valverde et al. 2002; Valverde y Schmidt 2000).

desenfocado, con extensas listas de objetivos por grado, es también típicamente un currículo con poca coherencia. Los estudios de los datos de TIMSS indican que procurar cubrir currículos extensos, da como resultado libros de texto y prácticas de enseñanza fragmentadas y sin continuidad. Es decir, los libros de texto y los docentes presentan objetivos de estas largas listas uno tras otro, en un intento de cubrirlos antes de que termine el año escolar con poco o ningún esfuerzo para explorar la relación entre estos conceptos o destrezas con ideas o temas fundamentales que los unen y dan significado. La pérdida de estas relaciones entre ideas (que forman parte de un riguroso aparato conceptual de las disciplinas) promueve que los estudiantes consideren las disciplinas como una serie de nociones aisladas y no como un todo disciplinario.

Estos estudios recientes también revelan importantes diferencias en cómo los sistemas educativos definen metas u objetivos. En un número de países de rendimientos promedio mediocres –podemos citar aquí el ejemplo de Estados Unidos– prevalece tradicionalmente una concepción en extremo estática de objetivos fundamentales. Esto es, los objetivos que se consideran fundamentales, se piensan que lo son a lo largo de toda la educación, lo que requiere que sean repetidos en muchos grados. La aritmética, por ejemplo, es una serie de contenidos y destrezas muy prominente a lo largo de todos los años de la enseñanza obligatoria. Incluso en octavo grado, cuando la mayoría de los países de alto rendimiento enfoca sus estándares en álgebra y geometría, la aritmética constituye una parte importante de los objetivos en Estados Unidos. Existen países en América Latina que presentan en sus políticas curriculares un perfil muy similar.

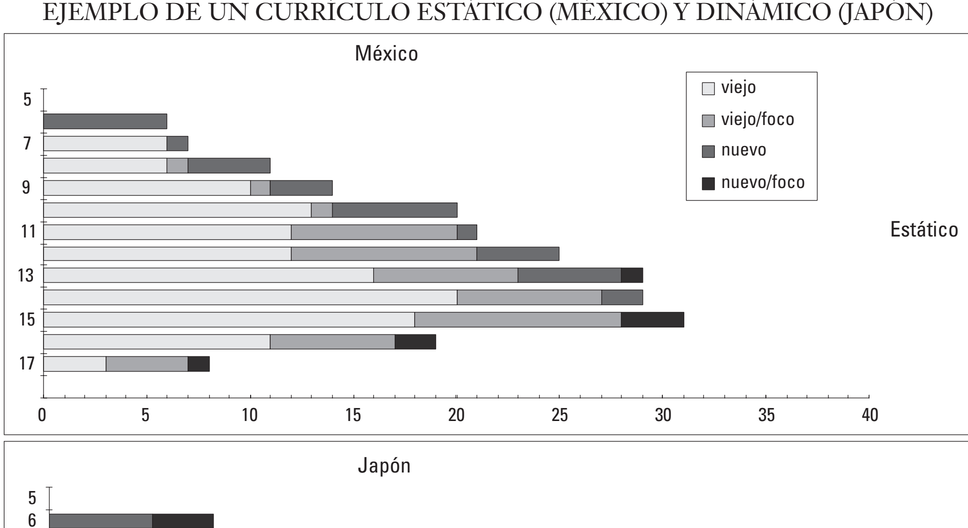
En países de alto rendimiento, cuando los objetivos son introducidos en el currículo, reciben una atención concentrada con la expectativa de que estos serán dominados de tal modo que los estudiantes están preparados para que se les presenten nuevos objetivos que deberán llegar a dominar en años subsiguientes. Metas enfocadas con claras prioridades constituyen los motores de una concepción dinámica de los estándares. En la mayoría de los países de alto rendimiento, cada grado ve la introducción de nuevos objetivos, los que reciben atención enfocada y que son explorados en profundidad como preparación para llegar a dominar objetivos cada vez más desafiantes en grados posteriores.

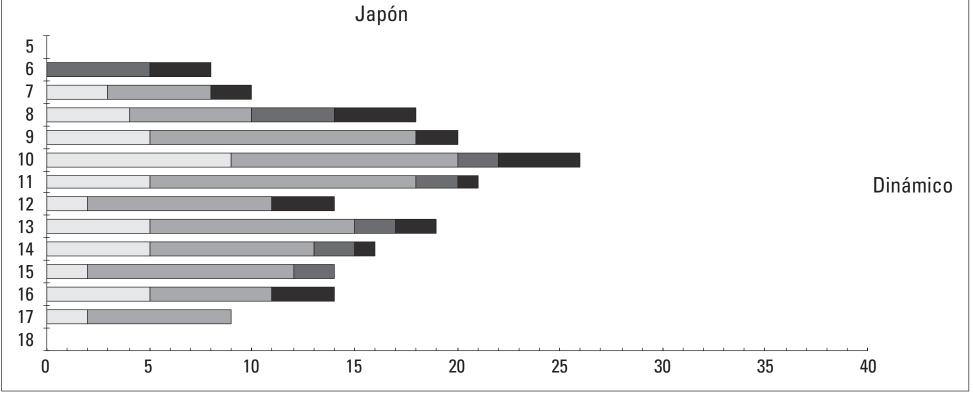
Una consecuencia de la falta de foco y coherencia, y de la perspectiva estática de los objetivos, es que tales currículos son poco rigurosos en comparación con otros países. Los materiales didácticos en estos países suelen cubrir una amplia gama de objetivos, la mayoría de ellos presentados por primera vez en la escuela primaria. El tratamiento es más bien ligero y superficial, no incluyendo mucho más que algoritmos y conocimientos de hechos e ideas aisladas. Estándares demandantes y propicios para estimular altos niveles de rendimiento parecen requerir contenidos y habilidades más sofisticados, explorados en

profundidad, a medida que los estudiantes avanzan en su escolarización. Los estándares rigurosos son el resultado de un proceso dinámico de transiciones hacia contenidos más complejos a partir del dominio cabal de objetivos más simples. El siguiente ejemplo proporciona un contraste entre el currículo estático de México, un país que escogió no publicar sus resultados durante su participación original en TIMSS, y el currículo dinámico de Japón, que estuvo entre los países de más alto rendimiento promedio en ese estudio.

*Gráfico Nº 1*

EJEMPLO DE UN CURRÍCULO ESTÁTICO (MÉXICO) Y DINÁMICO (JAPÓN)




Nota: cada columna horizontal representa el número de objetivos en términos de contenidos de Matemática que contiene el currículo prescrito para cada grado. El número de tópicos está representado en el eje horizontal. Los números en el eje vertical indican la edad normativa de los estudiantes en cada grado para propósitos de comparación internacional. La indicación «viejo» se usa para indicar objetivos que fueron introducidos en el currículo al menos en un grado previo. «Viejo/foco» indica objetivos propuestos en al menos un grado anterior –pero que se pretende que reciban atención concentrada en libros de texto y en la enseñanza. «Nuevo» se refiere a objetivos propuestos por primera vez, y «nuevo/foco» indica el subconjunto de nuevos objetivos que se pretende que reciban atención concentrada en el año en que son introducidos. Los perfiles de estos dos países representan concepciones distintas de la persecución de objetivos curriculares a lo largo de los años de escolarización.

Fuente: elaboración propia a partir de datos presentados por primera vez en Schmidt, McKnight, Valverde et al (1997).

Al estar ausentes discusiones acerca del rigor, y siendo escasos los esfuerzos específicos de contrastar la política curricular nacional con patrones internacionales entre los países de América Latina, es difícil tomar provecho de algunos de los más recientes hallazgos en el ámbito del estudio del currículo prescrito, a fin de aprovechar su potencial para propiciar oportunidades educativas que conduzcan a mayores niveles de aprendizaje. Estos mismos estudios han demostrado empíricamente cómo, en efecto, el currículo prescrito propicia y restringe oportunidades de aprendizaje y cómo estas a su vez afectan las probabilidades de lograr más altos niveles de rendimiento (McKnight y Valverde 1999; Schmidt et al. 2001).

## 2. Expectativas definidas a partir de un sistema de evaluación de altas implicaciones

Otra alternativa, que constituye también una de las primeras formas de definir expectativas que sigue vigente en muchos países del mundo, es mediante criterios de evaluación. Es decir, un sistema de pruebas con altas implicaciones[16] puede conformar el instrumento que formula las expectativas curriculares. Tal es el caso de las pruebas denominadas «Concourse» y «Baccalauréat» en Francia, y del «Abitur» alemán –así como en algunos países de la región latinoamericana donde ha habido, y siguen existiendo en algunos casos, pruebas de bachillerato de este tipo. En estos dos casos europeos, los criterios de evaluación fijan los contenidos de la formación inicial de docentes, su capacitación en servicio, el contenido de libros de texto, etc. De tal modo, el diseño de la prueba sirve para alinear distintos elementos de la política curricular y así enviar una clara señal de expectativas a docentes y estudiantes.

Generar expectativas claras a partir de un sistema de evaluación para la promoción tiene algunas importantes ventajas, puesto que la operacionalización del currículo (es decir, su formulación en términos mensurables) es requisito para generar instrumentos de medición, y al mismo tiempo los instrumentos en sí proporcionan ejemplos de los objetivos. Un ejemplo de una pregunta de biología en la prueba francesa de final de secundaria se presenta a continuación:

---

16 Por pruebas de «altas implicaciones» entendemos aquellas que se utilizan para tomar decisiones clave acerca de oportunidades de vida para los estudiantes. Entre estas pruebas incluimos aquellas que se usan para tomar decisiones acerca de la promoción de estudiantes de un grado o ciclo a otro, aquellas que asignan estudiantes a distintos tipos de oportunidades educativas (por ejemplo, a un programa vocacional o un programa preparativo para la educación superior), y las pruebas que determinan si un estudiante recibirá una credencial de conclusión de estudios secundarios. También consideramos de «altas implicaciones» aquellas pruebas que no tienen secuelas directas para el estudiante, pero sí para sus maestros o centro educativo, como aquellas que se vinculan a sistemas de incentivos y sanciones para docentes o escuelas.

RECUADRO N° 3
PARTE A (EXAMEN DE BACCALAURÉAT EN BIOLOGÍA DE PARIS, AMIENS, CRETEIL, LILLE, ROUEN, VERSAILLES[17])

«Se han tomado medidas de una muestra de células durante la interfase que precede a la mitosis. En ellas se ha descubierto lo siguiente:

- Se ha duplicado la cantidad de ADN en el núcleo, y
- Se ha incrementado el peso de las proteínas citoplasmáticas

Utilizando diagramas cuidadosamente seleccionados, concisos, y claramente anotados, ilustre los dos procesos que toman lugar en la preparación para la mitosis. Limite su contestación a las etapas que toman lugar en el núcleo de la célula. Sin detallar el proceso de síntesis de proteínas, explique cómo estas dos etapas preparan a la célula para conservar:

- su información genética, y
- su masa proteica de una generación a la siguiente.»

Evidentemente, la claridad de la visión en las pruebas en cuanto a objetivos del currículo de biología de la educación secundaria –tanto los contenidos, el nivel de expectativas en términos de rigor académico y expectativas acerca de lo que deben de saber hacer los estudiantes con el contenido aprendido–, en el caso de Francia, proporciona ejemplos utilísimos para la planificación de la instrucción. En el año 1993, cuando fue empleada esta pregunta en el Baccalauréat, el 51% de los estudiantes franceses rindieron esta prueba –todo aspirante a la educación superior debía rendirla (Britton y Raizen 1996).

Una limitante de la generación de expectativas claras a partir de un sistema de evaluación radica en su carácter deductivo, un tanto conductista, en el que se asume que las pruebas esclarecen los objetivos curriculares y que al responsabilizar a los maestros y estudiantes de su logro, se ha proporcionado información y motivación suficiente para cambiar prácticas pedagógicas y, por consiguiente, los logros de los estudiantes. Se admite entonces que el estímulo de la prueba basta para lograr el efecto de la renovación de la práctica pedagógica y el consiguiente mejoramiento en el logro. También se habla a veces de que fijar objetivos usando únicamente pruebas tenderá a restringir el currículo implementado a tal punto que lo único que se exploraría en el aula sería el contenido de las pruebas. Esto sin duda se transformaría en algo especialmente preocupante si las pruebas no evalúan contenidos y habilidades importantes.

[17] Ejemplo tomado de American Federation of Teachers and National Center for Improving Science Education (1994).

(El anterior ejemplo de Francia difícilmente puede llamarse ejemplo de una medición simple tendiente a bajar el nivel del currículo implementado).

Otro factor limitante es que no permite tomar en cuenta el aprendizaje o el progreso cuando se proponen metas absolutas para años terminales. Es decir, no es posible determinar distintos estadios de dominio de las metas que podrían mostrar los estudiantes conforme transitan por experiencias educativas sucesivas a lo largo de su carrera escolar.

## II. Evaluación

Quizá lo más evidente es la relativa corta historia de la evaluación estandarizada en América Latina. La gran mayoría de los sistemas nacionales de evaluación surgieron en la década de 1990. Por lo tanto, salvo excepciones notables como es el caso chileno, los distintos actores del sistema educativo se encuentran aun ahora familiarizándose con este tipo de evaluaciones. Muchos observadores y actores en los sistemas educativos de la región comentan que la introducción de la evaluación estandarizada ha sido precipitada, sin mucho trabajo preparatorio con las escuelas u otros actores. Inclusive se habla a veces de introducción forzosa o impuesta (por autoridades nacionales o por organismos internacionales como el Banco Mundial y el Banco Interamericano). No es por tanto sorprendente que algunas de las nociones que más se presentan en el temario de una discusión sobre la evaluación en Latinoamérica sean las de autoridad y poder. Como vimos anteriormente, son importantes temas en la tradición intelectual latinoamericana en el ámbito del currículo. Muchas de las objeciones a las pruebas estandarizadas son posiblemente atribuibles a la manera en que estas fueron introducidas y se usan en la región como instrumento de política educativa. En tal caso, la consideración de los méritos (o la falta de méritos) de estas evaluaciones no siempre deriva de un escrutinio de sus características técnicas, sino más bien proviene de conclusiones a partir de la percepción de sus orígenes y consecuencias políticas.

Las pruebas estandarizadas se introdujeron en un número importante de países como resultado de decisiones de autoridades, a veces inclusive como parte de convenios con donantes internacionales. Fue inusual que la introducción fuera precedida por esfuerzos de consulta nacional o con los actores educacionales en sentido más estricto. La introducción, además, no parece haber sido en muchos casos precedida por pilotos o esfuerzos experimentales destinados no solo al cuidadoso diseño de instrumentos válidos y apropiados, sino también como esfuerzo de experimentar cómo socializar la idea de las pruebas como instrumento de política, de demostrar sus beneficios y de encontrar cómo limitar sus consecuencias indeseables.

Esto fue seguido en muchos países por una difusión extensa de resultados, casi siempre con reacciones condenatorias a la educación (y a menudo en particular los docentes) por parte de la prensa y otros medios, inclusive a veces los propios gobiernos. Las publicaciones oficiales, y ciertamente las periodísticas, casi nunca se cuidaban (ni se cuidan en la actualidad) de reconocer las limitaciones de las pruebas y el carácter probabilístico de sus resultados (realizan inferencias acerca de muestras –en especial muestras de preguntas tomadas a partir del universo amplio de posibles preguntas que se pueden formular para determinar si una o un estudiante ha logrado dominar algún contenido escolar). De tal modo que los resultados conducían a pronunciamientos globales sobre la calidad de la educación nacional sin admisión de las importantes limitaciones e implicancias de tales generalizaciones.

De allí que no sorprenda la preocupación de muchos actores en los sistemas educativos por el tema del poder y la autoridad –la historia vivida en la región y la tradición intelectual hace que esa preocupación sea inevitable. Cuando se tiene además la oportunidad de examinar las pruebas o la documentación técnica que les corresponde, y se comprueba a veces sus debilidades técnicas– esto conduce a conclusiones casi inevitables. Dado lo anterior, al mismo tiempo que se observa la firmeza con que las autoridades hacen sus pronunciamientos acerca de lo que las pruebas revelan, se acentúa sin duda más la impresión de que la validez o calidad técnica no importa tanto como sí el uso de las pruebas para avanzar las agendas políticas de las autoridades del gobierno. Inclusive esta impresión seguramente es alimentada en aquellas ocasiones, que quizá no sean las más comunes pero sí muy conocidas, en que los ministerios o gobiernos no publican resultados. En esos casos los grandes operativos nacionales de medición, seguidos de un silencio absoluto por parte de las autoridades con respecto a los resultados, fácilmente lleva a la suspicacia de que la política tiene la primacía cuando de pruebas estandarizadas se trata. Estas pruebas casi siempre comparten características técnicas básicas –son a menudo instrumentos con preguntas de opción o escogencia múltiple y en principio casi todas referidas a normas, y por tanto muchos asocian ese tipo de prueba con la agenda en política educativa de las autoridades gubernamentales.

Semejante a lo anterior, también pesa que la introducción de las pruebas estandarizadas en muchos países se dio casi en forma simultánea con reformas curriculares que enfocaban nuevas perspectivas pedagógicas. Estas perspectivas, a menudo basadas en nociones de la construcción psicosocial del conocimiento, el reconocimiento de la existencia de múltiples tipos de inteligencia, etc., casi nunca ha sido encarnada en las pruebas estandarizadas. La disonancia entre los nuevos currículos propuestos y las perspectivas curriculares de las pruebas estandarizadas puede ser el origen de muchas posiciones críticas respecto de las pruebas. Por ejemplo, podría explicar la gran cantidad de opiniones de que las

pruebas no pueden responder a las necesidades de información de docentes o expertos en currículo. También, la opinión de que los supuestos básicos de las técnicas de medición cuantitativa en educación son incompatibles con las nuevas perspectivas curriculares parece encontrar sustento en esto.

Todo lo anterior pareciera en algunos casos alimentar la idea de que la evaluación estandarizada es un producto de las necesidades de información de las autoridades gubernamentales. Por tanto es una evaluación incompatible con las perspectivas pedagógicas-didácticas promovidas en los nuevos currículos y que las agencias de evaluación están poco preocupadas por la validez de sus inferencias acerca de la calidad de la educación.

Esta percepción de la evaluación estandarizada señalada en párrafos anteriores pareciera confluir naturalmente con lo que podríamos llamar una suerte de «nihilismo métrico». ¿En qué consiste este «nihilismo métrico»? Se trata de una posición que se observa con frecuencia entre muchos críticos de la evaluación en América Latina, que considera imposible que la medición cuantitativa arroje informaciones que iluminen el fenómeno de la educación. Una variante de esta posición es una convicción de la invalidez sustancial de la medición cuantitativa, aunque mitigada a veces por un reconocimiento de que las pruebas estandarizadas pueden recoger alguna información superficial, insuficiente para el propósito de entender el sentido profundo del aprendizaje que quizá requieran algunos de sus actores clave (como por ejemplo docentes), pero que satisfacen las necesidades superficiales de información de diseñadores de políticas e interesados que están bastante alejados del trabajo educativo en aulas.

Frente a estas formas de comprender la evaluación de una manera tan reduccionista, simplificadora y limitante, existen otras posiciones que se pueden tomar al considerar que las necesidades de información de agentes externos al aula son legítimas y que pueden ser suplidas por sistemas de pruebas. El público en muchos países de América Latina ha demostrado que demanda informaciones, como por ejemplo: ¿cómo se compara lo que logra mi hija o hijo en la escuela con lo que debiera lograr allí de acuerdo al currículo oficial?, ¿cómo se compara la calidad de distintas escuelas, o la calidad de la educación de mi provincia con respecto a otras escuelas o provincias? La educación es una importante inversión social y resulta difícil rechazar la idea de que los inversionistas tienen legítimo interés en monitorear algunos indicadores acerca de cuán bien va su inversión En esta posición, que podríamos llamar de «postempiricismo métrico», se considera fundamental responder a estas necesidades de información asumiendo la responsabilidad de proporcionar indicadores educativos con grados aceptables de validez y confiabilidad, y al mismo tiempo procurando desarrollar instrumentos de evaluación capaces de rendir apropiada cuenta de la inversión social, sin vulnerar inapropiadamente las instituciones educativas. Esta posición reconoce la legitimidad de muchas de las críticas a la evaluación y

en paralelo busca fundamentar las evaluaciones respecto del reconocimiento riguroso del carácter probabilístico y perfectible de la medición de logros y aprendizajes, en donde se reconoce que la validez y la confiabilidad son cuestiones de grado que exigen el continuo mejoramiento y refinamiento técnico.

Lamentablemente, el nihilismo métrico implica que las escuelas estarían libres de los llamados a la transparencia y apertura al escrutinio ciudadano al que están sujetas otras instituciones públicas o instituciones privadas de interés público. Una posición postempiricista constituye una sólida y viable alternativa, porque puede sobreponerse al empiricismo ingenuo de algunos defensores de la evaluación que no enfrentan de lleno su carácter probabilístico, perfectible, y que no enjuician con suficiente criticidad los problemas que supone la habitual ausencia de consideraciones rigurosas de los problemas de validez y confiabilidad que a menudo aquejan a los sistemas de evaluación actuales en la región. Esto es importante porque los desafíos técnico-políticos que enfrentan los sistemas de medición de la calidad en América Latina son considerables.

El compromiso de los gobiernos con la evaluación en la región es a menudo tibio, sujeto a los vaivenes de la inconsistencia de prioridades políticas. Existen, por ejemplo, unidades a cargo de las pruebas en algunos países que en muchas oportunidades no tienen claridad con respecto a su presupuesto, inclusive a menos de seis meses antes de aplicar una prueba a nivel nacional. Estas unidades a veces deben contratar personal, desarrollar y validar pruebas, y aplicarlas a nivel nacional, todo en cuestión de unos pocos meses. Incluso algunas veces las unidades no cuentan con los recursos necesarios para cubrir funciones vitales: recientemente un país de la región que aplicó pruebas en múltiples grados y asignaturas a nivel sistémico por primera vez no contaba con los recursos asignados para la captura y análisis de los datos, encontrándose las pruebas almacenadas en una bodega sin un plan claro para afrontar la necesidad. El problema fue resuelto acudiendo a la ayuda internacional. Este es un ejemplo de la fragilidad que a menudo caracteriza la evaluación en América Latina. Una fragilidad que dificulta el desarrollo técnico y la consolidación política de la evaluación.

La vaguedad y falta de priorización en los currículos prescritos de la región también son un importante obstáculo para la evaluación. Muy a menudo las pruebas nacionales no están alineadas con el currículo prescrito, precisamente porque no existen criterios claros con los cuales lograr convergencia. Con frecuencia un enfoque curricular inspira el currículo nacional, y un enfoque diferente fundamenta la prueba nacional, contribuyendo a una cacofonía de mensajes que difícilmente pueden ayudar a impulsar enseñanzas y aprendizajes significativos. Este es un obstáculo especialmente prominente en países federativos como Argentina y Brasil, donde o no han existido currículos nacionales, pero sí pruebas nacionales, o los currículos nacionales son débiles o poco comparables con los currículos estatales o provinciales. Es esperanzador ver los

casos de Chile, que ha incluido la evaluación y el currículo en una misma unidad ministerial, y de Honduras y Guatemala que en sus nuevos esfuerzos por escribir estándares han también abordado el desarrollo de sistemas de evaluación que miden con referencia a estos estándares y que buscan las herramientas técnicas para hacerlo cada vez con mayor rigor. En Chile, sin embargo, en la recién aprobada Ley General de Educación (2009), se crea un nuevo marco institucional en el que la evaluación es separada del ministerio y asumida por una Agencia de Calidad de la Educación externa. La alineación de currículo y evaluación es crucial para la validez de ambos, pero la región aún se encuentra experimentando con respecto a las disposiciones institucionales apropiadas para estas dos áreas. En la actualidad, asumir la alineación de currículo y evaluación como un desafío institucional prioritario no es común en la región.[18]

## III. Una estructura de políticas en estándares y evaluación

La sociedad y los Estados en América Latina enfrentan importantes retos en la gestión del sistema educativo, los que al considerarse inevitablemente conducirán al replanteamiento de sus sistemas de políticas en evaluación y currículo. Uno de los retos más importantes, y el más relevante para consideraciones en esta área de políticas educativas, es el tema de la calidad. La preocupación por la calidad resulta ser primordialmente una preocupación por el contenido de la escolarización. Es decir, es una preocupación por lo que se aprende en la escuela y, por consiguiente, por lo que se enseña. Asumir responsabilidad por lo que se aprende en la escuela constituye un desafío importante para la política educativa, un desafío que, como hemos visto, muchos países han enfrentado con distintos niveles de éxito.

---

[18] Por otro lado, Chile también confronta importantes problemas en su sistema de evaluación: está desafiado por una larga tradición de publicaciones de tablas comparativas o «rankings» de sus escuelas en la prensa que usan los resultados de las pruebas nacionales. Se interpretan estos como indicadores de efectividad o calidad de las escuelas sin que el instrumento presente las características apropiadas para tales inferencias. Desde 1999, el SIMCE presenta las características técnicas para ser comparable de año en año. Son muy pocos los sistemas de evaluación en América Latina que presentan esta característica. Aun cuando el diseño tradicional del SIMCE se ha llamado «referido al currículo», presenta esencialmente las características (lo que es común en la región) de una prueba referida a normas. Las preguntas (o ítems) se escriben con referencia a una tabla de especificaciones cuyas categorías principales se derivan del currículo, pero todos los procesos de selección, depuración y desarrollo subsiguientes de los ítems se hacen siguiendo los procedimientos clásicos de las pruebas referidas a normas. Desde 2006, el SIMCE de 4º básico se reporta de acuerdo con criterios definidos de logro, y el proceso de desarrollo de Mapas de Progreso como parte de la política curricular proporciona una herramienta de gran valor para pasar a un modelo de medición más acorde a una evaluación referida a criterios curriculares.

El desafío mayor que hemos identificado en este capítulo es el de contar con un sistema de mensajes convergentes. Por tal razón, hemos hecho hincapié también en lo importante que es que las especificaciones de metas curriculares sean formuladas en forma clara, operacional y ejemplificada. Las políticas curriculares en los programas nacionales de estudios, estándares, o currículo nacional son a menudo vagas, ambiguas, dispersas y poco demandantes en relación a otras partes del mundo. Las pruebas estandarizadas que se administran en los mismos países casi nunca se desarrollan para usar el currículo prescrito de manera rigurosa como referente en la definición de áreas a evaluar, puntajes de corte, niveles de rendimiento, etc. Ante este desafío, un sistema articulado de políticas en evaluación y currículo surge como una alternativa prometedora y responsable. Un sistema de instrumentos políticos de este tipo podría resolver los problemas de falta de claridad, de especificidad y de debilidad en la convergencia de mensajes. Asimismo, podría estructurar un debate acerca de las políticas educativas apropiadas a las aspiraciones de competitividad y excelencia de nuestras naciones, tendiente a resolver los problemas que suponen objetivos magros.

En este capítulo hemos concentrado nuestra atención en evaluación y estándares, que son los temas que nos atañen. Sin embargo, es importante recordar que estos no pueden ser más que un subconjunto de los instrumentos necesarios para buscar calidad en un sistema educativo. Sin atender los otros temas que forman parte de lo que se aborda en este libro (financiamiento, formación del magisterio, etc.), ningún sistema de evaluación y estándares será capaz de garantizar calidad. En este capítulo tratamos el tema de la calidad a partir de políticas de este tipo, reconociendo que deben formar parte de un conjunto comprensivo de instrumentos que buscan complementarse en la búsqueda de la calidad.

Un sistema articulado de políticas en evaluación y currículo debe contar, como punto de partida, con una definición clara de lo que se entenderá por «calidad» en el logro escolar. Partiendo de esta definición fundamental, debidamente apoyada en instrumentos de política curricular pertinentes, se puede diseñar un sistema de instrumentos de monitoreo y evaluación capaz de proporcionar indicadores de cómo los agentes y distintos componentes institucionales del sistema educativo persiguen las metas de calidad, y sobre cuáles características sistémicas (como prácticas de instrucción, formación inicial y continua de docentes, libros de texto, etc.) son mejores para lograr estas metas. Un sistema articulado que siga estos principios será capaz de producir evidencias para la toma de decisiones por parte de las autoridades educativas y también ayudará a orientar la atención de quienes toman decisiones políticas y de la sociedad cuando se trate de considerar diferentes alternativas de política y gestión para impulsar metas de calidad educativa.

El fin que se persigue con la articulación de un sistema como el descrito, es el de establecer mecanismos permanentes, sólidamente institucionalizados, para

establecer y revisar metas de aprendizaje, y de medición de su logro. Asimismo, se puede buscar fundamentar un marco institucional de políticas y normas con respecto al uso que se quiere y puede hacer de los resultados de evaluaciones, que sirva a su vez como piedra angular de un conjunto de políticas claras en donde se establecen y se aceptan responsabilidades de todos los actores que intervienen o deben de intervenir en el proceso de alcanzar las metas de calidad del sistema educativo. Sobre este tipo de políticas de responsabilización (*accountability*) se está debatiendo en gran manera en América Latina. Un sistema articulado de instrumentos de política curricular, como el que revisaremos a continuación, puede servir como base de un sistema de responsabilización formal, con su conjunto de incentivos y/o sanciones específicas para diversos actores y su sistema de pruebas censales. También puede ser útil para racionalizar las políticas de currículo, sin formar parte de un sistema formal que asigna responsabilidades recíprocas a agencias y a actores individuales.

Un sistema articulado de políticas de currículo y evaluación contesta mediante instrumentos políticos claros las expectativas educativas del sistema en torno a tres preguntas fundamentales –preguntas que formulan la gran mayoría de los actores y de las instituciones que forman parte del sistema educativo. Estas preguntas son, en orden:

- ¿Qué deben aprender los estudiantes?
- ¿Qué deben enseñar los maestros?
- ¿Cómo evaluar y monitorear lo que se aprende y se enseña?
- ¿Qué acciones se deben de tomar de acuerdo con los resultados de la evaluación y monitoreo?[19]

Un sistema articulado formula, sucesivamente, instrumentos que juntos constituyen los seis elementos o funciones básicas del sistema –no siempre se formulan como instrumentos distintos, pero las seis funciones son comunes a la mayoría de los casos de establecimiento de sistemas articulados de políticas de currículo y evaluación basadas en criterios de calidad que se detallan a continuación.

---

[19] Frente a estas preguntas de la política curricular y la evaluación existen otras relativas a las condiciones en que se aprende y se enseña, que comúnmente competen a otros aspectos de la política educativa. Por ejemplo, determinar qué deben aprender estudiantes y enseñar profesores implica compromisos con respecto a la dotación física de ambientes escolares, entrenamiento de docentes, libros de texto, etc. La idea de políticas articuladas de currículo y evaluación apunta a que pueden también orientar otras áreas de la política educativa, como por ejemplo el financiamiento. Es poco frecuente que las políticas de financiación consideren lo que parece ser la pregunta fundamental: «¿cuánto cuesta una educación de calidad?». Un sistema de políticas articuladas en el currículo prescrito como el que describimos aquí puede proporcionar insumos para contestar esa pregunta.

*Tabla Nº 1*

FUNCIONES BÁSICAS DE UN SISTEMA ARTICULADO DE POLÍTICAS DE CURRÍCULO Y EVALUACIÓN

| *Elemento* | *Función* | *Ámbito técnico / político* |
|---|---|---|
| 1. Marco curricular | Define el ámbito curricular a partir del cual se tomarán decisiones acerca de metas y prioridades. | Corresponde al currículo 'potencial' –el universo curricular del cual las metas y prioridades serán una muestra. |
| 2. Metas de aprendizaje | Establecen los elementos curriculares de los que se hará cargo el sistema educativo. | Definen el currículo prescrito. |
| 5. Estándares de aprendizaje | Esclarecen los objetivos de aprendizaje mediante su operacionalización y ejemplificación. | |
| 4. Hitos de logro | Esclarecen en forma operacional y ejemplificada los puntos críticos de progreso hacia el logro de las expectativas de aprendizaje. | |
| 5. Expectativas pedagógicas | Esclarecen, en forma operacional y ejemplificada, las oportunidades de aprendizaje que demandan las expectativas. | Establecen elementos para la implementación del currículo intencional. |
| 6. Especificaciones de monitoreo y evaluación. | Establecen sobre cuáles expectativas e hitos el sistema de evaluación se responsabilizará por documentar. | Establecen las responsabilidades de monitoreo. |

Fuente: elaboración propia.

*Resumen de elementos/funciones en el establecimiento de claras expectativas acerca de lo que se aprenderá en la escuela*

*Marco curricular.* Las metas específicas son parte de un intento por guiar el sistema educativo de acuerdo con concepciones de calidad de resultados. En la mayoría de los países, los recursos para la educación no bastan para abordar todas las metas posibles –ni tan siquiera todas las deseables. Las metas se deben de escoger como las más deseables dentro del rango de lo posible. Un marco curricular de base es un documento que describe el rango completo de las posibilidades en cada disciplina, aunque probablemente no todas estas posibilidades se concreten entre las expectativas que se fijen subsiguientemente. Los marcos establecen no solo contenidos disciplinarios (por ejemplo, tópicos en Matemática), sino que además qué se espera que los estudiantes sean capaces de hacer a partir de esos contenidos. También pueden establecer temas generales y orientaciones básicas que tienen una relación menos directa con contenidos y habilidades estrictamente disciplinarios (por ejemplo, la importancia de la Matemática para una sociedad).

*Metas de aprendizaje.* Las metas son acuerdos acerca de los fines deseados de la escolarización, tomados a partir del rango de metas posibles. Los marcos establecen lo posible, las metas establecen que lo que se ha acordado es lo más

deseable y lo prioritario. Mientras que los marcos son inclusivos, estableciendo un rango amplio de posibilidades, las metas son exclusivas, excluyendo todo salvo algunas metas escogidas. Existen dos requisitos básicos para tener metas: se necesita establecer fines deseables para la escolarización y se necesitan acuerdos. Sin una visión clara de los objetivos es imposible establecer metas, pero sin acuerdos o consensos acerca de esos objetivos solo resultarán metas en sentido formal, sin mucha posibilidad de servir como herramientas para el cambio. Es vital tener presente, a la hora de tomar decisiones acerca de metas, que un modelo de enseñanza-aprendizaje que dé prioridad a la construcción psicosocial y a la aplicación del conocimiento, al desarrollo de actitudes, y a procesos centrados en el estudiante, requiere muchos más recursos que modelos educativos anteriores. Uno de los recursos más importantes es el tiempo: un modelo como el descrito simplemente tomará más tiempo para explorar a profundidad una cantidad menor de contenidos o temas. De tal modo, se deben tomar decisiones muy precisas acerca de cuáles elementos curriculares son prioritarios y cuáles no. Esto subraya la importancia de las formulaciones de metas.

*Estándares de aprendizaje*. A diferencia de las metas, los estándares se formulan en términos específicos, objetivos y mensurables[20] u observables. Son el resultado de un proceso tanto político como técnico-pedagógico de búsqueda de consensos. Esclarecen el contenido de los acuerdos acerca de las metas (más generales de aprendizaje) y por ello a menudo descubren desacuerdos o diferencias entre distintos actores sociales que no se evidencian en el lenguaje general de esas metas. La función de los estándares es, por un lado, dar solidez a los acuerdos acerca de metas y, por el otro, proporcionar un mensaje curricular básico para el desarrollo de planes de estudio, libros de texto, etc. Ambas funciones se cumplen al proporcionar especificaciones claras y sin ambigüedad en las metas.

*Hitos*, benchmarks *o indicadores progresivos de logro.* Hemos establecido que los estándares esclarecen los fines u objetivos de la escolarización en términos de las áreas curriculares. Describen la meta al final de una serie de procesos y actividades de cambio a lo largo de la trayectoria escolar. Estos cambios ocurren a lo largo del tiempo, y deben ser monitoreados con regularidad para determinar si distintos elementos del sistema se acercan a los fines esperados.

[20] Por mensurable se entiende en forma en que se proporcionan elementos para reconocer cuando el estudiante ha cumplido la expectativa, es decir, por mensurable entendemos «susceptible de ser confirmado empíricamente» y por lo tanto no estamos pretendiendo circunscribir la noción a únicamente aquello que pueda ser medido en las pruebas escritas, o de opción múltiple, sino un rango amplio de procedimientos de verificación empírica tanto cuantitativos como cualitativos.

Los hitos o indicadores de logros son señales en el camino que nos permiten determinar el progreso de los alumnos hacia el logro de la meta. En muchos países, las funciones de estándares e indicadores de logro se cumplen en un mismo instrumento.

*Expectativas pedagógicas.* Los estándares y los hitos de logro deben ser interpretados para el quehacer diario de la instrucción, y dos de los intérpretes clave son el magisterio y los productores de material didáctico. A ambos les compete el desafío de traducir los estándares en términos de oportunidades educativas concretas. Un sistema de gestión del sistema educativo basado en el esclarecimiento de expectativas no puede considerarse completo con la sola publicación de los acuerdos. El acuerdo o consenso en las expectativas debe incluir un acuerdo o compromiso de actuar de acuerdo a las decisiones, formulando y ejecutando programas concretos de desarrollo de material didáctico, y de formación y capacitación docente alineados con los estándares.

*Especificaciones de monitoreo y evaluación*. Estas especificaciones representan acuerdos acerca de cómo se medirá el cambio y progreso educativo en diversos aspectos del logro académico, oportunidades de aprendizaje y otros elementos del sistema educativo. Las especificaciones de monitoreo y evaluación describen las mediciones que se tomarán para determinar si los indicadores de logro se han alcanzado. Consiste en una especificación detallada de cómo procederá el sistema en la recolección y valoración sistemática de información acerca de la necesidad (demanda), el diseño, la implementación, el impacto y la efectividad de políticas, programas y otras prácticas políticas y administrativas del sistema educativo. Tales prácticas de evaluación y monitoreo estarían diseñadas para informar de un modo tendiente a mejorar las propias prácticas evaluadas. La evaluación en un sistema de políticas curriculares, basado en expectativas claras de aprendizaje, tiene el propósito fundamental de monitorear, evaluar y reportar públicamente el progreso del ministerio, otras instancias administrativas, directores de centros educativos, docentes, estudiantes, etc. hacia el logro de expectativas y otros fines establecidos para el sistema. Corresponden, por tanto, a una definición de las responsabilidades de cada actor frente al sistema educativo y a la sociedad en general –define los elementos acerca de los cuales rendirán cuentas estudiantes, maestros, escuelas, Ministerio de Educación, etc.

Un ejemplo útil de algunos de los aspectos que discutimos en este esquema, es el caso de Australia Occidental, en donde se prescriben estándares de aprendizaje que al mismo tiempo incorporan hitos progresivos de logro y que se muestra a continuación.[21]

[21] Este ejemplo se ha tomado de Ferrer (2006).

RECUADRO Nº 4

## AUSTRALIA OCCIDENTAL

### Estándares y niveles de desempeño para las Ciencias Naturales
### Eje: Energía y Transformación

| Nivel 1 | Nivel 2 | Nivel 3 | Niveles 4 a 8 |
|---|---|---|---|
| Los estudiantes comprenden que la energía es necesaria para diferentes propósitos en la vida cotidiana. | Los estudiantes comprenden las diferentes maneras en que se transfiere la energía, y también que ésta es utilizada en sus distintas formas con diferentes propósitos. | Los estudiantes comprenden diferentes patrones de uso y algunos tipos de transferencia de energía. | |
| *En este nivel:* Los estudiantes observan y describen experiencias concretas a partir de la exploración sensorial del mundo físico que los rodea. Reconocen algunas fuentes de energía, sus usos y reconocen que puede cambiar. | *En este nivel:* Los estudiantes describen las causas y efectos relacionados con fenómenos observables y reconocen las transformaciones ocurridas en un plano real y concreto. | *En este nivel:* Los estudiantes realizan generalizaciones basadas en observaciones concretas del mundo físico. Comprenden que la energía puede ser transferida y transformada, y describen sus propias observaciones. Establecen relaciones entre patrones de uso, fuentes y receptores de la energía. Utilizan estos patrones para realizar predicciones sobre fenómenos y relaciones que trascienden el entorno inmediato y su experiencia personal. | |
| *Especificaciones y ejemplos* Los alumnos reconocen que la energía puede ser transferida o transformada. Describen cómo la alimentación les provee de energía para moverse y jugar. Reconocen que la energía de una batería se transfiere a un juguete eléctrico para hacerlo mover, encender una luz o producir un sonido. Observan que empujar o tirar de un auto de juguete es una forma de energía para hacer que se desplace. | *Especificaciones y ejemplos* Los alumnos comprenden que la energía puede ser transformada para utilizarse con diferentes propósitos. Reconocen que encender un calentador eléctrico o cocina produce calor y que, mientras uno se utiliza para calentar un ambiente, el otro sirve para cocinar alimentos. Comprenden que la energía puede transferirse de un objeto a otro. Describen cómo la electricidad, el gas o la madera utilizados en una cocina pueden calentar una sartén y luego los alimentos que contiene. Identifican situaciones en que se transfiere energía, tales como el pedaleo en una bicicleta o el viento en la propulsión de un velero. | *Especificaciones y ejemplos* Los alumnos relacionan la transferencia de energía con el uso que se le da a esa energía. Describen que los objetos en movimiento –como un bate–, contienen energía que puede ser transferida a otros (una pelota). Comienzan a utilizar terminología específica como energía cinética para identificar energía y movimiento. Explican que la energía puede ser transferida a través de acciones como empujar, jalar o golpear, tales como la dirección y cantidad de energía utilizada para lograr un gol en fútbol. Comprenden que parte de la energía de los neumáticos de un auto | |

Un sistema de políticas educativas basado en expectativas claras de aprendizaje requiere de una operacionalización curricular de la calidad para que puedan diseñarse indicadores que permitan el monitoreo sistémico, es decir, la observación de cómo los diversos elementos, actores y agencias del sistema educativo trabajan para conseguir la calidad en los aprendizajes de los y las estudiantes. La definición de expectativas sirve para delimitar los aspectos del currículo con el fin de generar indicadores y vincularlos entre sí. Estos proporcionan un «lenguaje operacional» para ligar la definición curricular de la calidad con los indicadores del sistema. Un sistema bien definido de expectativas con respecto a lo que se debe aprender y enseñar, se usa para caracterizar indicadores de logro, aprendizaje, cobertura curricular, oportunidades para aprender, etc. También puede proporcionar criterios para evaluar libros de textos, prácticas pedagógicas, material didáctico, etc. según los mismos términos de referencia. Contando con un conjunto de especificaciones operacionales de la calidad como el descrito, el sistema será capaz de generar datos estructurados para monitorear la calidad de forma cada vez más válida y útil.

## Conclusión: una mirada al futuro

En este capítulo se ha procurado revisar la experiencia latinoamericana en cuanto a los estándares y la evaluación, contrastándola con patrones mundiales. Se ha buscado retratar el significado de estos contrastes y se han presentado ideas de cómo algunos de los retos más prominentes se podrían resolver.

Lo anterior deja muchas preguntas que parecen fundamentales, tanto para sistemas nacionales de política curricular y evaluación como para la comunidad educativa y sociedad latinoamericana en general:

- ¿Es acaso posible o deseable reconducir los esfuerzos de evaluación, reconociendo falencias en su introducción, procurando compensar estas con esfuerzos mayores de socialización y genuina consulta?
- ¿Es posible aprender de estas falencias en la introducción de la evaluación de tal modo que los nacientes esfuerzos por desarrollar estándares las eviten, procurando así no desatar preocupaciones relacionadas con la autoridad y el control derivados también de la insuficiente socialización y consulta?
- ¿Cómo hacer transparente los procesos de diseño de estándares, aplicación, análisis y difusión de resultados de pruebas de tal modo que los destinatarios estén más persuadidos de que existe una genuina preocupación por la calidad de los estándares, y la precisión, validez y confiabilidad de las informaciones generadas por las pruebas?

- ¿Es deseable o no incluir mayor preocupación por el rigor y la comparabilidad mundial de las políticas curriculares en los esfuerzos de reforma curricular o diseño de estándares en América Latina?
- ¿Será posible, con el desarrollo de los sistemas de evaluación en América Latina, lograr que la política curricular se fundamente más en la evidencia empírica acerca de lo que logran o no los y las estudiantes en sus escuelas como complemento de sus fundamentos teóricos?
- ¿Cómo procurar que las preocupaciones por precisión, validez y confiabilidad precedan esfuerzos por utilizar los resultados de pruebas estandarizadas en la toma de decisiones, o inclusive generar pronunciamientos generales acerca de la calidad de la educación?
- ¿Será importante aumentar pronto la cantidad de herramientas de medición de los sistemas de mediciones, dando prioridad al desarrollo de técnicas y procedimientos mejor alineados con las prioridades curriculares e instruccionales en la región?
- ¿Cómo separar la discusión acerca del papel político de las evaluaciones de la discusión de sus méritos y falencias técnicas, de modo que se reconozca que ambas discusiones son muy importantes y en parte convergentes, pero que también involucran distintos criterios y objetivos?

La creciente discusión regional acerca de la evaluación y las reacciones ante la publicación de resultados de pruebas a nivel de las distintas naciones del área, sugieren que estas preguntas deben preocupar a autoridades, la comunidad educativa y la sociedad en general, si se ha de avanzar en el desarrollo de la frecuentemente nombrada «cultura de evaluación».

Es claro, a partir de la investigación empírica internacional, que el interés que se ha tenido en la política curricular por parte de muchas naciones es válido. Puede ser un instrumento clave para propiciar el acceso equitativo a experiencias educativas ricas y significativas. De ahí que América Latina debe considerar cuánto de lo que se ha aprendido a nivel internacional es aplicable a sus esperanzas de mejorar la calidad de sus sistemas educativos.

## Referencias

American Federation of Teachers and National Center for Improving Science Education (1994). *What College-Bound Students Abroad Are Expected to Know About Biology: Exams from England and Wales, France, Germany and Japan. Plus a Comparative Look at the United States*. Vol. 1, *Defining World Class Standards*. Washington, D.C.: American Federation of Teachers.

Benavot, A. (1992). «Curricular Content, Educational Expansion and Economic Growth.» *Comparative Education Review* 36 (2):150-174.

Benavot, A., Cha, Yun-Kyung, Kamens, D., Meyer J. W., and Suk-Ying Wong (1991). «Knowledge for the Masses: World Models and National Curricula, 1920–1986.» *American Sociological Review* 56 (February):85-100.

Bernstein, B. (1973). *Class, Codes, and Control*. Vol. 2. London: Routledge & Kegan Paul.

Britton, E. D., and S. A. Raizen (eds.) (1996). *Examining the Examinations: An International Comparison of Science and Mathematics Examinations for College-Bound Students*. Dordrecht, The Netherlands: The National Center for Improving Science Education/Kluwer Academic Publishers.

Cardoso, F. H., and E. Faletto (1980). *Dependency and Development in Latin America*. Berkeley: University of California Press.

Devon, R. F. (1983). «The Uneasy Triangle: Theory, Policy, and Reality in Educational Resource Flows Between Center and Periphery. *Studies in Comparative International Development* 18 (1-2):113-138.

Dirección Nacional de Educación Inicial y Primaria (2000). Estructura Curricular Básica de Educación Primaria de Menores: Programa Curricular de Primer Ciclo de Educación Primaria de Menores. Lima: Ministerio de Educación de Perú.

Ferrer, G. (2006). Estándares de currículo: algunas tendencias internacionales e implicancias para su implementación en América Latina. Lima: PREAL-GTEE.

——. 2006. *Sistemas de evaluación de los aprendizajes en América Latina: balance y desafíos*. Santiago, Chile: Grupo de Trabajo sobre Evaluación y Estándares-PREAL.

Ferrer, J. G., Valverde, G. A. y J. M. Esquivel Alfaro (2006). «Aspectos del currículo prescrito en América Latina: revisión de tendencias contemporáneas en currículo, indicadores de logro, estándares y otros instrumentos. En P. Arregui (ed.), *Sobre estándares y evaluaciones en América Latina*. Lima: Grupo de Trabajo Sobre Estándares y Evaluación, PREAL.

Levy, F., and R. J. Murnane (2005). *The New Division of Labor: How Computers are Creating the Next Job Market*. Princeton, N.J.: Princeton University Press.

McKnight, C. C., and G. A. Valverde (1999). «Explaining TIMSS Mathematics Achievement: A Preliminary Survey.» In G. Kaiser, E. Luna and I. Huntley (eds.), *International Comparisons in Mathematics Education*. London: Falmer Press.

Ministerio de Educación de Perú (1997). *Logros básicos e indicadores*. Lima: Ministerio de Educación de Perú.

MONBUSHO-Ministry of Education, Science and Culture (1995). Guide for Elementary School. Tokyo: National Institute for Educational Research.

Ravela, P., Wolfe, R., Valverde, G., and J. M. Esquivel (2001). *Los próximos pasos: ¿cómo avanzar en la evaluación de aprendizajes en América Latina?* Santiago de Chile: Grupo de Trabajo sobre Evaluación y Estándares-PREAL.

Schmidt, W. H., McKnight, C. C., Cogan, L. S., Jakwerth, P. M., and R. T. Houang (1999). *Facing the Consequences: Using TIMSS for a Closer Look at U.S. Mathematics and Science Education*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, HsingChi, Wiley, D. E., Cogan, L.S., and R. G. Wolfe (2001). *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*. San Francisco: Jossey-Bass Publishers.

Schmidt, W. H., McKnight, C. C., Raizen, S. A.,. Jakwerth, P. M., Valverde, G. A., Wolfe, R. G., Britton, E. D., Bianchi, L. J., and R. T. Houang (1997). *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., and D. E. Wiley (1997). *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Mathematics*. Vol. 1. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt, W. H., Raizen, S. A., Britton, E. D., Bianchi, L. J., and R. G. Wolfe (1997). *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Science*. Vol. 2. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schneider, H. (2005). «Rapid ICT Change and Workplace Knowledge Obsolescence: Causes and Proposed Solutions.» *Research Publication*. Cambridge, Mass.: The Berkman Center for Internet and Society at Harvard Law School.

Secretaría de Estado de Educación, Bellas Artes y Cultos (1995). *Plan Decenal de Educación en Acción Transformación Curricular en Marcha: Nivel Básico*. Santo Domingo: SEEBAC.

Sica, A., and H. Prechel (1981). «National Political-Economic Dependency in the Global Economy and Educational Development.» *Comparative Education Review*.

Stigler, J. (1997). Videotape Classroom Study. http://nces.ed.gov/timss/video/index.html: National Center for Education Statistics [Online].

Torraco, R. J. (2002). «Cognitive Demands of New Technologies and the Implications for Learning Theory.» *Human Resource Development Review* 1 (4):439-467.

Valverde, G. A. (2000). «Strategic Themes in Curriculum Policy Documents: An Exploration of TIMSS Curriculum Analysis Data.» *International Journal of Educational Policy Research and Practice.* 1 (2):133-152.

——. (1997). «Evaluación y estándares de currículo en una era de reforma educativa. En B. Álvarez y M. Ruiz-Casares (eds.), *Evaluación y reforma educativa: opciones de política.* Washington, D.C.: U.S. Agency for International Development.

Valverde, G. A., Bianchi, L. J., Houang, R. T., Schmidt, W. H., and R. G. Wolfe (2002). *According to the Book: Using TIMSS to Investigate theTtranslation from Policy to Pedagogy in the World of Textbooks*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Valverde, G. A., and W. H. Schmidt (2000). «Greater Expectations: Learning from other Nations in the Quest for World-class Standards in US School Mathematics and Science.» *Journal of Curriculum Studies* 32 (5):651-687.

Westbury, I., and Chao-Hseng Hsu (1996). «Structures of Curriculum Governance and Classroom Practice in Mathematics.» *Educational Evaluation and Policy Analysis* 18 (2):123-140.